

RECEIVED  
CENTRAL FAX CENTER

FEB 20 2008

HEWLETT-PACKARD COMPANY  
Intellectual Property Administration  
P.O. Box 272400  
Fort Collins, Colorado 80527-2400

PATENT APPLICATION

ATTORNEY DOCKET NO. 10018003-1IN THE  
UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor(s): Jian Fan

Confirmation No.: 9516

Application No.: 10/044,558

Examiner: Brian Q. Le

Filing Date: January 11, 2002

Group Art Unit: 2524

Title: TEXT EXTRACTION AND ITS APPLICATION TO COMPOUND DOCUMENT IMAGE COMPRESSION

Mail Stop Appeal Brief-Patents  
Commissioner For Patents  
PO Box 1450  
Alexandria, VA 22313-1450

## TRANSMITTAL OF APPEAL BRIEF

Transmitted herewith is the Appeal Brief in this application with respect to the Notice of Appeal filed on Dec. 20, 2007.☒ The fee for filing this Appeal Brief is \$510.00 (37 CFR 41.20).☐ No Additional Fee Required.

(complete (a) or (b) as applicable)

The proceedings herein are for a patent application and the provisions of 37 CFR 1.138(a) apply.

☐ (a) Applicant petitions for an extension of time under 37 CFR 1.136 (fees: 37 CFR 1.17(a)-(d)) for the total number of months checked below:☐ 1st Month  
\$120☐ 2nd Month  
\$460☐ 3rd Month  
\$1050☐ 4th Month  
\$1640☐ The extension fee has already been filed in this application.☐ (b) Applicant believes that no extension of time is required. However, this conditional petition is being made to provide for the possibility that applicant has inadvertently overlooked the need for a petition and fee for extension of time.Please charge to Deposit Account 08-2025 the sum of \$ 510. At any time during the pendency of this application, please charge any fees required or credit any over payment to Deposit Account 08-2025 pursuant to 37 CFR 1.25. Additionally please charge any fees to Deposit Account 08-2025 under 37 CFR 1.18 through 1.21 inclusive, and any other sections in Title 37 of the Code of Federal Regulations that may regulate fees.☒ A duplicate copy of this transmittal letter is enclosed.☐ I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to:  
Commissioner for Patents, Alexandria, VA 22313-1450  
Date of Deposit:

OR

☒ I hereby certify that this paper is being transmitted to the Patent and Trademark Office facsimile number (571)273-8300.

Date of facsimile: February 19, 2008

Typed Name: Judy H. Chung

Signature: Judy H. Chung

Respectfully submitted,

Jian Fan

By: Ashok K. Mannava

Ashok K. Mannava

Attorney/Agent for Applicant(s)

Reg No.: 45,301

Date: February 20, 2008

Telephone: (703) 652-3822

(Rev 10/07)(ApB/Brief)

Total number of pages: 27

**RECEIVED**  
**CENTRAL FAX CENTER****FEB 20 2008**

Attorney Docket No.: 10018003-1

**HEWLETT-PACKARD COMPANY**  
Intellectual Property Administration  
P.O. Box 272400  
Fort Collins, Colorado 80527-2400**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

**Inventor(s):** Jian Fan **Confirmation No.:** 9516  
**Serial No.:** 10/044,558 **Examiner:** Brian Q. Le  
**Filed:** January 11, 2002 **Group Art Unit:** 2624  
**Title:** TEXT EXTRACTION AND ITS APPLICATION TO COMPOUND DOCUMENT  
IMAGE COMPRESSION

**MAIL STOP APPEAL BRIEF - PATENTS**  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

**APPEAL BRIEF - PATENTS**

Sir:

This is an Appeal Brief in connection with the decisions of the Examiner in a final Office Action mailed September 25, 2007 and in connection with the Notice of Appeal mailed December 20, 2007. It is respectfully submitted that the present application has been more than twice rejected. Each of the topics required in an Appeal Brief and a Table of Contents are presented herewith and labeled appropriately.

02/22/2008 SSITHIB1 00000041 082025 10044558

01 FC:1402 510.00 DA

**PATENT**

Atty Docket No.: 10018003-1

App. Scr. No.: 10/044,558

**TABLE OF CONTENTS**

(1) Real Party In Interest .....	3
(2) Related Appeals And Interferences .....	3
(3) Status Of Claims .....	3
(4) Status of Amendments .....	4
(5) Summary Of Claimed Subject Matter .....	5
(6) Grounds of Rejection to be Reviewed on Appeal .....	8
(7) Arguments .....	9
A. The rejection of claims 1-3, 6, and 23-26 under 35 USC § 102(b) is improper because Lee fails to teach each and every element of claims 1-3, 6, and 23-26. ....	9
B. The rejection of claim 5 under 35 USC § 103(a) is improper because Lee in view of Kodaira fails to teach or suggest all the features of claim 5. ....	13
C. The rejection of claims 8, 10, and 15-17 under 35 USC § 103(a) is improper because Lee in view of Hashimoto fails to teach or suggest all the features of claims 8, 10, and 15-17. ....	14
(8) Conclusion .....	15
(9) Claim Appendix .....	16
(10) Evidence Appendix .....	24
(11) Related Proceedings Appendix .....	25

**PATENT**

Atty Docket No.: 10018003-1

App. Ser. No.: 10/044,558

**(1) Real Party In Interest**

The real party in interest is Hewlett-Packard Development Company, L.P.

**(2) Related Appeals And Interferences**

There are no other appeals or interferences related to this case.

**(3) Status Of Claims**

Claims 1-3, 5-17, 20 and 22-26 are pending of which claims 1, 23 and 26 are independent.

Claims 4, 18, 19, 21, 27, and 28 are canceled.

All pending claims 1-3, 5-17, 20 and 22-26 are hereby appealed.

Claims 1-3, 6, and 23-26 stand rejected under 35 USC § 102(b) as allegedly being anticipated by Lee et al. U.S. Patent No. 5,583,659 ("Lee").

Claim 5 stands rejected under 35 USC § 103(a) as allegedly being unpatentable over the combination of Lee in view of Kodaira et al. U.S. 6,868,183 ("Kodaira").

Claims 8, 10, and 15-17 stand rejected under 35 USC § 103(a) as allegedly being unpatentable over the combination of Lee and Hashimoto et al. U.S. 6,987,045 ("Hashimoto")

Claims 7, 9, 11-14, 20, and 22 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

**PATENT**

Atty Docket No.: 10018003-1

App. Ser. No.: 10/044,558

**(4) Status of Amendments**

No amendment was filed subsequent to the final Office Action dated April 2, 2007.

**PATENT**

Atty Docket No.: 10018003-1

App. Ser. No.: 10/044,558

**(5) Summary of Claimed Subject Matter**

Text extraction is a common process in optical character recognition (OCR) systems. When pixels that represent text are identified in OCR systems, they can be separated from other pixels and compressed more efficiently. Current techniques to identify pixels, such as global and adaptive thresholding, are used with documents having simple text printed on a white background. However, the use of different colors and styles of text, as well as colored and graphical backgrounds has become increasingly more common (paragraphs 3-11). The embodiments described herein provide more accurate systems and methods for identifying and extracting pixels that represent text.

For example, claim 1 recites a method to identify text-like pixels from an image (FIG. 1), the method comprising:

classifying a plurality of individual pixels within a mask within the image as either edge or non-edge, wherein a pixel (i,j) is located at the center of the mask (paragraphs 29-31 and FIG. 4);

determining whether the pixel (i,j) is an edge pixel or a non-edge pixel (paragraphs 29-31);

determining whether pixels having connectivity with the pixel (i,j) are edge pixels or non-edge pixels (paragraphs 29-31); and

performing edge-bounded averaging to determine line segments, wherein the edge-bounded averaging includes finding one of either

**PATENT**

Atty Docket No.: 10018003-1

App. Ser. No.: 10/044,558

an average value of only the edge pixels having connectivity with pixel (i,j), in response to determining that pixel (i,j) is an edge pixel or

an average value of only the non-edge pixels having connectivity with pixel (i,j) in response to determining that pixel (i,j) is a non-edge pixel (paragraphs 29-31).

Claim 23 recites a system for identifying text-like pixels from an image (FIG. 12), the system comprising:

a processor for classifying a plurality of individual pixels within a mask within the image as either edge or non-edge (paragraphs 83 and 84), wherein a pixel (i,j) is located at the center of the mask; determining whether the pixel (i,j) is an edge pixel or a non-edge pixel; determining whether pixels having connectivity with the pixel (i,j) are edge pixels or non-edge pixels; and performing edge-bounded averaging to determine line segments (paragraphs 29-31), the edge-bounded averaging including one of either

finding an average value of only the edge pixels having connectivity with pixel (i,j), in response to determining that pixel (i,j) is an edge pixel or

finding an average value of only the non-edge pixels having connectivity with pixel (i,j) in response to determining that pixel (i,j) is a non-edge pixel (paragraphs 29-31).

Claim 26 recites a computer readable storage medium on which is embedded one or more computer programs comprising a set of instructions that when executed by a processing circuit performs a method of processing a digital image (FIG. 12, paragraphs 83 and 84), the method comprising:

**PATENT**

Atty Docket No.: 10018003-1

App. Ser. No.: 10/044,558

classifying a plurality of individual pixels within a mask within the digital image as either edge or non-edge, wherein a pixel (i,j) is located at the center of the mask (paragraphs 29-31);

determining whether the pixel (i,j) is an edge pixel or a non-edge pixel (paragraphs 29-31);

determining whether pixels having connectivity with the pixel (i,j) are edge pixels or non-edge pixels (paragraphs 29-31); and

performing edge-bounded averaging to determine line segments, wherein the edge-bounded averaging includes finding one of either

an average value of only the edge pixels having connectivity with pixel (i,j), in response to determining that pixel (i,j) is an edge pixel or

an average value of only the non-edge pixels having connectivity with pixel (i,j) in response to determining that pixel (i,j) is a non-edge pixel (paragraphs 29-31).



**PATENT**

Atty Docket No.: 10018003-1  
App. Ser. No.: 10/044,558

**(6) Grounds of Rejection to be Reviewed on Appeal**

A. Whether claims 1-3, 6, and 23-26 should have been rejected under 35 USC § 102(b) as being anticipated by Lee.

B. Whether claim 5 should have been rejected under 35 USC § 103(a) as being unpatenable over the combination of Lee in view of Kodaira.

C. Whether claims 8, 10, and 15-17 should have been rejected under 35 USC § 103(a) as being unpatenable over the combination of Lee and Hashimoto.

**PATENT**

Atty Docket No.: 10018003-1

App. Ser. No.: 10/044,558

**(7) Arguments****A. The rejection of claims 1-3, 6, and 23-26 under 35 USC § 102(b) is improper because Lee fails to teach each and every element of claims 1-3, 6, and 23-26.**

The test for determining if a reference anticipates a claim, for purposes of a rejection under 35 U.S.C. § 102, is whether the reference discloses all the elements of the claimed combination, or the mechanical equivalents thereof functioning in substantially the same way to produce substantially the same results. As noted by the Court of Appeals for the Federal Circuit in *Lindemann Maschinenfabrick GmbH v. American Hoist and Derrick Co.*, 221 USPQ 481, 485 (Fed. Cir. 1984), in evaluating the sufficiency of an anticipation rejection under 35 U.S.C. § 102, the Court stated:

Anticipation requires the presence in a single prior art reference disclosure of each and every element of the claimed invention, arranged as in the claim.

Therefore, if the cited reference does not disclose each and every element of the claimed invention, then the cited reference fails to anticipate the claimed invention and, thus, the claimed invention is distinguishable over the cited reference. This rejection is respectfully traversed because Lee fails to disclose all the features of independent claims 1, 23, and 26, and the claims that depend therefrom.

**1. Lee fails to teach "edge-bounded averaging," as recited in claims 1, 23, and 26.**

**PATENT**

Atty Docket No.: 10018003-1  
App. Ser. No.: 10/044,558

Independent claims 1, 23, and 26 recite that "edge-bounded averaging includes finding an average value of only edge pixels having connectivity with pixel (i,j), in response to determining that pixel (i,j) is an edge pixel or an average value of only non-edge pixels having connectivity with pixel (i,j) in response to determining that pixel (i,j) is a non-edge pixel." In contrast to the claimed features, Lee discloses averaging all pixels within a mask. Lee does not teach averaging either only edge pixels or only non-edge pixels having connectivity based on the classification of the center pixel. In other words, Lee averages every pixel within a mask (or window), whereas the claims recite averaging "only" a particular subset of pixels, i.e. edge or non-edge having connectivity.

The Final Office Action alleges that Lee teaches these features in column 8, lines 34-43. Lines 34 and 35 of the cited passage state that "if area gradient...exceeds the threshold value, GT, then pixel (i,j) lies in the vicinity of an edge." The Examiner erroneously concludes that this sentence teaches determining if a pixel is an edge pixel or a non-edge pixel. The Examiner then appears to rely on the phrase "in this case," in line 36 of the cited passage, as an implication that the additional steps explained in the passage of ascertaining  $L_{max}$  and  $L_{min}$  and calculating  $L_{avg}$  are only performed on edge pixels. The Final Office Action also seems to place particular emphasis on the sentence which reads "thereafter, method 200 advances to block 248 which calculates an average pixel intensity value." However, it is not clear how these extracts from Lee purportedly teach the claimed features.

Regarding the Examiner's erroneous conclusion that Lee teaches edge-bounded averaging, the cited passage of Lee, like the rest of the Lee reference and the other cited art, fails

**PATENT**

Atty Docket No.: 10018003-1

App. Ser. No.: 10/044,558

to teach, or even suggest, performing edge-bounded averaging with only edge pixels or only non-edge pixels, as recited in independent claims 1, 23, and 26. In contrast to the claimed features, Lee teaches averaging all pixels within a window without determining if the non-center pixels in the window are edge or non-edge and also without considering the connectivity of the pixels in the window. For example, the Applicant notes column 7, lines 15-35, in which Lee provides a detailed explanation of averaging pixel windows. Here, Lee refers to the 3x3 pixel window depicted in FIG. 3A and details a calculation in which every pixel in the window is averaged without regard to the connectivity of the pixels or a classification of the other pixels in the window as either edge or non-edge.

Moreover, in column 6, lines 50-55, Lee further teaches that "each and every" pixel goes through the same averaging process, which is "identical across all such pixels." Therefore, Lee clearly teaches averaging all pixels in a window and not just the connected only edge or only non-edge pixels, as recited in the independent claims. Lee provides absolutely no suggestion that only edge or only non-edge pixels can, or should, be averaged. Similarly, Lee makes no mention of averaging connected pixels. This is because, Lee teaches averaging all pixels within a pixel window, as opposed to averaging "only edge pixels" or "only non-edge pixels" having connectivity with the center pixels. As such, Lee fails to teach or suggest the above-recited features of independent claims 1, 23, and 26.

2. Lee fails to teach "determining whether the pixels having connectivity with pixel (i,j) are edge pixels or non-edge pixels," as recited in claims 1, 23, and 26.

**PATENT**

Atty Docket No.: 10018003-1

App. Ser. No.: 10/044,558

Lee also fails to teach "determining whether the pixels having connectivity with pixel (i,j) are edge pixels or non-edge pixels." Connectivity is described by way of example in paragraph 31 of the originally filed specification and illustrated in Figure 3. In contrast to this claimed feature, Lee discloses classifying only the center pixel (i,j) within a pixel window as edge or non-edge. Lee then repeats this process for other center pixels (i,j). Lee does not teach or suggest determining any relationships, such as connectivity, between pixels within individual windows. Lee simply averages all pixels within each window without regard to their connectivity. Therefore, Lee fails to teach determining if pixels having connectivity to pixel (i,j) are edge pixels or non-edge pixels.

**3. Lee fails to teach performing edge-bounded averaging to determine "line segments," as recited in claims 1, 23, and 26.**

In addition, Lee fails to teach performing edge-bounded averaging to determine line segments. Initially, as set forth above, Lee fails to teach performing the claimed edge-bounded averaging. However, Lee also fails to teach or suggest any averaging process to determine line segments. Lee is drawn to thresholding an image to determine if an image is a character or background. Lee fails to teach or suggest performing averaging to determine a line segment.

Accordingly, it is respectfully submitted that the rejection is improper because Lee fails to teach each and every element of independent claims 1, 23, and 26. Therefore, withdrawal of this rejection and allowance of the claims is respectfully requested.

PATENT

Atty Docket No.: 10018003-1

App. Ser. No.: 10/044,558

**B. The rejection of claim 5 under 35 USC § 103(a) is improper because Lee in view of Kodaira fails to teach or suggest all the features of claim 5.**

The test for determining if a claim is rendered obvious by one or more references for purposes of a rejection under 35 U.S.C. § 103 is set forth in MPEP § 706.02(j):

To establish a *prima facie* case of obviousness, three basic criteria must be met. First, there must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings. Second, there must be a reasonable expectation of success. Finally, the prior art reference (or references when combined) must teach or suggest all the claim limitations. The teaching or suggestion to make the claimed combination and the reasonable expectation of success must both be found in the prior art and not based on applicant's disclosure. *In re Vaeck*, 947 F.2d 488, 20 USPQ2d 1438 (Fed. Cir. 1991).

Therefore, if the above-identified criteria are not met, then the cited reference(s) fails to render obvious the claimed invention and, thus, the claimed invention is distinguishable over the cited reference(s).

Lee and Kodaira, taken singly or in combination, fail to teach or suggest the elements of independent claim 1, from which claim 5 depends, for the reasons set forth above. Accordingly, claim 5 is allowable at least by virtue of its dependence on claim 1 and withdrawal of this rejection and allowance of the claims is respectfully requested.

**PATENT**

Atty Docket No.: 10018003-1

App. Ser. No.: 10/044,558

**C. The rejection of claims 8, 10, and 15-17 under 35 USC § 103(a) is improper because Lee in view of Hashimoto fails to teach or suggest all the features of claims 8, 10, and 15-17.**

Lee and Hashimoto, taken singly or in combination, fail to teach or suggest the elements of independent claim 1, from which claims 8, 10, and 15-17 depend, for the reasons set forth above. Accordingly, claims 8, 10, and 15-17 are allowable at least by virtue of their dependence on claim 1 and withdrawal of this rejection and allowance of the claims is respectfully requested.

RECEIVED  
CENTRAL FAX CENTER

FEB 20 2008

## PATENT

Atty Docket No.: 10018003-1

App. Ser. No.: 10/044,558

## (8) Conclusion

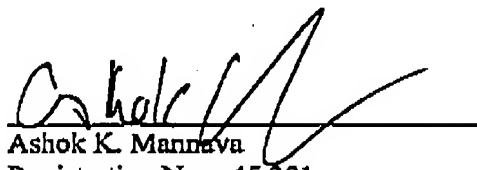
For at least the reasons given above, the rejection of claims 1, 3-29, 31-33 and 35 is improper. Accordingly, it is respectfully requested that such a rejection by the examiner be reversed and these claims be allowed. Attached below for the Board's convenience is an Appendix of claims 1, 3-29, 31-33 and 35 as currently pending and on appeal.

Please grant any required extensions of time and charge any fees due in connection with this Appeal Brief to deposit account no. 08-2025.

Respectfully submitted,

Dated: February 20, 2008

By

  
Ashok K. Mannava  
Registration No.: 45,301

MANNAVA & KANG, P.C.  
1240 Waples Mill Road  
Suite 300  
Fairfax, VA 22030  
(703) 652-3822  
(703) 865-5150 (facsimile)



**PATENT**

Atty Docket No.: 10018003-1

App. Ser. No.: 10/044,558

**(9) Claim Appendix****CLAIMS:**

1. A method to identify text-like pixels from an image, the method comprising:

classifying a plurality of individual pixels within a mask within the image as either edge or non-edge, wherein a pixel (i,j) is located at the center of the mask;

determining whether the pixel (i,j) is an edge pixel or a non-edge pixel;

determining whether pixels having connectivity with the pixel (i,j) are edge pixels or non-edge pixels; and

performing edge-bounded averaging to determine line segments, wherein the edge-bounded averaging includes finding one of either

an average value of only the edge pixels having connectivity with pixel (i,j), in response to determining that pixel (i,j) is an edge pixel or

an average value of only the non-edge pixels having connectivity with pixel (i,j) in response to determining that pixel (i,j) is a non-edge pixel.

2. The method of claim 1, further comprising:

examining sub-blobs of pixels within the image; and

performing sub-blob connectivity analysis.

3. The method of claim 2, further comprising:

**PATENT**

Atty Docket No.: 10018003-1

App. Ser. No.: 10/044,558

identifying and classifying edges of pixels within the image;  
 performing filling to further classify pixels within the image;  
 performing consistency analysis of pixels within the image;  
 performing pixel connectivity analysis of pixels within the image; and  
 identifying text pixels within the image.

4. (Canceled).

5. The method of claim 1, further comprising performing color space conversion of the image.

6. The method of claim 1, further comprising smoothing the image.

7. The method of claim 1, wherein a Gaussian lowpass filter is applied to the image, the filter being

$$f_{i,j} = ke^{-\alpha^2 \left[ (i-c)^2 + (j-c)^2 \right] / c^2}$$

where k is a normalizing factor such that  $\sum_{i,j} f_{i,j} = 1.0$ , c is the center of the filter and  $\alpha = 1.0$

8. The method of claim 3, wherein the step of identifying and classifying edges of pixels within the image further comprises,

classifying every pixel as NON EDGE, WHITE EDGE or BLACK EDGE.

**PATENT**

Atty Docket No.: 10018003-1

App. Ser. No.: 10/044,558

9. The method of claim 8, wherein the step of identifying and classifying edges of pixels within the image further comprises:

(1) calculating a vertical gradient  $G'_{i,j}$ , a horizontal gradient  $G''_{i,j}$ , and the magnitude of gradient  $M_{i,j}$  using the formula,

$$G'_{i,j} = (y_{i+1,j-1} + 2y_{i+1,j} + y_{i+1,j+1}) - (y_{i-1,j-1} + 2y_{i-1,j} + y_{i-1,j+1})$$

$$G''_{i,j} = (y_{i+1,j-1} + 2y_{i+1,j} + y_{i+1,j+1}) - (y_{i-1,j-1} + 2y_{i-1,j} + y_{i-1,j+1})$$

$$M_{i,j} = \sqrt{(G'_{i,j})^2 + (G''_{i,j})^2}$$

Where  $y_{i,j}$  is a pixel luminance value at an index  $i,j$

(2) calculating a discrete Laplacian (a second directive):

$$L_{i,j} = (y_{i-2,j} + y_{i+2,j} + y_{i,j-2} + y_{i,j+2}) - 4y_{i,j}$$

(3) classifying every pixel as the following:

If  $M_{i,j} > T_e$  then

If  $L_{i,j} < 0$

Classify pixel at  $(i,j)$  as WHITE EDGE

Else

Classify pixel at  $(i,j)$  as BLACK EDGE

Endif

Else

**PATENT**

Atty Docket No.: 10018003-1

App. Ser. No.: 10/044,558

Classify pixel at  $(i,j)$  as NON EDGE

Endif.

10. The method of claim 1, wherein the step of performing edge-bounded averaging further comprises:

starting from a first side of a line proceeding to a second side of the line identifying consecutive segments of pixels as NON EDGE, WHITE EDGE or BLACK EDGE.

11. The method of claim 1, wherein the step of performing edge-bounded averaging comprises:

computing the edge-bounded averaging for at least eight locations including both end points of a central interior, both end points of a left edge segment, both end points of a right edge segment, a right end point of a left interior and a left end point of a right interior.

12. The method of claim 11, further comprising:

classifying the central interior as NON TEXT, BLACK INTERIOR or WHITE INTERIOR based upon the edge-bounded averaging values.

13. The method of claim 3, wherein the step of performing filling to further classify pixels within the image comprises:

classifying segments as NON TEXT; and

**PATENT**

Atty Docket No.: 10018003-1

App. Ser. No.: 10/044,558

examining segments classified as NON TEXT to determine whether they may be reclassified as BLACK INTERIOR, BLACK EDGE, WHITE INTERIOR or WHITE EDGE.

14. The method of claim 3, wherein the step of performing vertical consistency analysis of pixels within the image comprises:

examining pixels not yet classified as NON TEXT to determine whether they are BLACK INTERIOR, BLACK EDGE, WHITE INTERIOR or WHITE EDGE.

15. The method of claim 3, wherein the step of performing pixel connectivity analysis of pixels within the image comprises:

identifying aggregates of pixels having been identified as candidates for text, the aggregates being sub-blobs; and

collecting statistics with respect to each sub-blob, wherein said statistics are selected from the group consisting of total number of pixels, sums of color values, number of border pixels, number of broken border pixels and horizontal run length.

16. The method of claim 2, wherein the step of examining sub-blobs of pixels within the image comprises:

examining each sub-blob to determine whether it is NON TEXT.

17. The method of claim 3, wherein the step of identifying text pixels comprises:

**PATENT**

Atty Docket No.: 10018003-1

App. Ser. No.: 10/044,558

examining each sub-blob to classify each pixel as either a text pixel or a non-text pixel.

Claims 18-19 (Canceled).

20. The computer readable storage medium according to claim 26, said one or more computer programs further comprising a set of instructions for:

performing pixel connectivity analysis of pixels within the digital image identifying aggregates of pixels having been identified as candidates for text, the aggregates being sub-blobs;

collecting each sub-blobs statistics: total number of pixels, sums of color values, number of border pixels, number of broken border pixels and horizontal run length; and

counting sums of each luminance and chroma.

21. (Canceled).

22. The computer readable storage medium according to claim 26, said one or more computer programs further comprising a set of instructions for:

performing pixel connectivity analysis of pixels within the digital image by identifying aggregates of pixels having been identified as candidates for text, the aggregates being sub-blobs;

**PATENT**

Atty Docket No.: 10018003-1

App. Ser. No.: 10/044,558

collecting each sub-blobs statistics: total number of pixels, sums of color values, number of border pixels, number of broken border pixels and horizontal run length; and counting sums of each  $Y, C_r, C_b$ .

23. A system for identifying text-like pixels from an image, the system comprising:

a processor for classifying a plurality of individual pixels within a mask within the image as either edge or non-edge, wherein a pixel (i,j) is located at the center of the mask; determining whether the pixel (i,j) is an edge pixel or a non-edge pixel; determining whether pixels having connectivity with the pixel (i,j) are edge pixels or non-edge pixels; and performing edge-bounded averaging to determine line segments, the edge-bounded averaging including one of either

finding an average value of only the edge pixels having connectivity with pixel (i,j), in response to determining that pixel (i,j) is an edge pixel or

finding an average value of only the non-edge pixels having connectivity with pixel (i,j) in response to determining that pixel (i,j) is a non-edge pixel.

24. The system of claim 23, wherein the processor also examines sub-blobs of pixels within the image; and performs sub-blob connectivity analysis.

25. The system of claim 24, wherein the processor also identifies and classifies edges of pixels within the image; performs vertical filling to further classify pixels within the image; performs

**PATENT**

Atty Docket No.: 10018003-1

App. Ser. No.: 10/044,558

vertical consistency analysis of pixels within the image; performs pixel connectivity analysis of pixels within the image; and identifies text pixels.

26. A computer readable storage medium on which is embedded one or more computer programs comprising a set of instructions that when executed by a processing circuit performs a method of processing a digital image, the method comprising:

classifying a plurality of individual pixels within a mask within the digital image as either edge or non-edge, wherein a pixel (i,j) is located at the center of the mask;

determining whether the pixel (i,j) is an edge pixel or a non-edge pixel;

determining whether pixels having connectivity with the pixel (i,j) are edge pixels or non-edge pixels; and

performing edge-bounded averaging to determine line segments, wherein the edge-bounded averaging includes finding one of either

an average value of only the edge pixels having connectivity with pixel (i,j), in response to determining that pixel (i,j) is an edge pixel or

an average value of only the non-edge pixels having connectivity with pixel (i,j) in response to determining that pixel (i,j) is a non-edge pixel.

Claims 27-28 (Canceled).



**PATENT**

**Atty Docket No.: 10018003-1**

**App. Ser. No.: 10/044,558**

**(10) Evidence Appendix**

None.

**PATENT**

**Atty Docket No.: 10018003-1**

**App. Ser. No.: 10/044,558**

**(11) Related Proceedings Appendix**

None.